

# Mapper

```
class MAPPER
```

```
  method INITIALIZE()
```

```
    Create empty temp_list
```

```
    Set old_user_id to some non-positive integer (say 0)
```

```
  method MAP(line l)
```

```
    info <- split the line
```

```
    user_id <- info[0]
```

```
    movie_id <- info[1]
```

```
    rating <- info[2]
```

```
    if rating >= 4.0 do
```

```
      if old_user_id != user_id do
```

```
        sort temp_list
```

```
        for all movie_i in temp_list do
```

```
          for all movie_j > movie_i in temp_list do
```

```
            EMIT(movie_i, movie_j, 1)
```

```
        temp_list <- []
```

```
        old_user_id <- user_id
```

```
    else
```

```
      temp_list.append(movie_id)
```

# Reducer

**class** REDUCER

**method** INITIALIZE()

Read movies.csv file and store the movie names into a global dictionary

where movie\_id is key and movie\_name is the value

set threshold to any value (say 1000)

**method** REDUCE(movie\_i, movie\_j, counts [c1, c2, .....])

sum <- 0

**for all** c in counts **do**

sum <- sum + c

**if** sum > threshold **do**

EMIT(movie\_i, movie\_j, sum)